IN THE SPECIFICATION

Please amend the specification as follows:

Page 4, lines 5 - 33:

Preferably, the determining step includes setting the singing-starting time point of the first phoneme to a time point earlier than the ~~signing-starting~~ singing-starting time point represented by the time information.

According to this singing voice-synthesizing method, the phonetic unit information, the time information, and the singing length information are inputted in timing earlier than the singing-starting time point represented by the time information, and a phonetic unit transition time length is formed based on the phonetic unit information. Further, a singing-starting time point and a singing duration time of the first phoneme and a singing-starting time point and a singing duration time of the second phoneme are determined based on the generated phonetic unit transition time length. As a result, as to the first and second phonemes, it is possible to determine desired ~~signing-starting~~ singing-starting time points before or after the singing-starting time point represented by the time information, or determine singing duration times different from the singing length represented by the singing length information, whereby natural ~~signing~~ singing sounds can be produced as the first and second singing phonetic units. For example, if the singing-starting time point of the first phoneme can be set to a time point earlier than the singing-starting time point represented by the time information, it is possible to make the rise of a consonant sufficiently earlier than the rise of a vowel to thereby synthesize singing voices close to human singing voices.

Page 6, lines 24-33:

This singing voice-synthesizing apparatus implements the ~~signing~~ singing sound-

synthesizing method according to the first aspect of the invention, and hence the same advantageous effects described as to this method can be obtained. Further, since the apparatus is configured such that the phonetic unit transition time length is read from the storage section, the construction of the apparatus or the processing executed thereby can be simple even if the number of ~~signing~~ singing phonetic units is increased.

Page 8, lines 3-11:

This singing voice-synthesizing apparatus implements the ~~signing~~ singing sound-synthesizing method according to the first aspect of the invention, and hence the same advantageous effects described as to this method can be obtained. Further, since the apparatus is configured such that the phonetic unit transition time length is read from the storage section, the construction of the apparatus or the processing executed thereby can be simple even if the number of ~~signing~~ singing phonetic units is increased.

Page 8, line 27 – page 9, line 6:

To attain the above object, according to a fifth aspect of the invention, there is provided a ~~signing~~ singing sound-synthesizing apparatus comprising an input section that inputs phonetic unit information representative of a phonetic unit, time information representative of a singing-starting time point, singing length information representative of a singing length, and effects-imparting information, for a singing phonetic unit, and a singing voice-synthesizing section that generates a singing voice formed by the phonetic unit, based on the phonetic unit information, the time information, and the singing length information which have been inputted by the input section, the singing voice synthesizing section imparting effects to the singing voice based on the

effects-imparting information inputted by the input section.

Page 9, line 31 – page 10, line 24:

To attain the above object, according to a sixth aspect of the invention, there is provided a singing voice-synthesizing apparatus comprising an input section that inputs phonetic unit information representative of a phonetic unit, time information representative of a singing-starting time point, and singing length information representative of a singing length, for a singing phonetic unit, in timing earlier than the ~~signing-starting~~ singing-starting time point, a setting section that randomly sets a new singing-starting time point, within a predetermined time range extending before and after the singing-starting time point, based on the time information inputted by the input section, and a singing voice-synthesizing section that generates a singing voice formed by the phonetic unit, based on the phonetic unit information and the singing length information which have been inputted by the input section, and the singing-starting time point set by the setting section, the singing voice synthesizing section starting generation of the ~~signing~~ singing sound at the new singing-starting time point set by the setting section.

According to this singing voice-synthesizing apparatus, a new singing-starting time point is randomly set within a predetermined time range extending before and after the singing-starting time point represented by the time information, and a singing voice is generated at the set singing-starting time point. This makes it possible to synthesize more natural singing voices with variations in ~~signing-starting~~ singing-starting timing.

Page 11, line 23:

FIG. 8 is a diagram showing information stored in a vibrato DB;

Page 13, lines 15-16:

FIG. 32 is a flowchart showing a vowel singing length-calculating process;

Page 20, lines 2-7:

FIG. 4 shows information contained in the performance data. The performance data contains performance information necessary for singing one syllable, and the performance information contains note information, phonetic unit track information, transition ~~truck~~ track information, and vibrato track information.

Page 23, lines 20-35:

(b) "Sil_C" represents a phonetic unit transition from silence to a consonant ~~constant~~, and, for example, in FIG. 6B, corresponds to a combination of the preceding phonetic unit "Sil" and the following consonant "s", not shown.

(c) "C_V" represents a phonetic unit transition from a consonant ~~constant~~ to a vowel, and, for example, in FIG. 6B, corresponds to a combination of the preceding consonant "s", not shown, and the following vowel "a", not shown.

(d) "Sil_V" represents a phonetic unit transition from silence to a vowel, and, for example, in FIG. 6B, corresponds to a combination of the preceding phonetic unit "Sil" and the following vowel "a".

(e) "pV_C" represents a phonetic unit transition from a preceding vowel to a consonant ~~constant~~, and, for example, in FIG. 6B, corresponds to a combination of the preceding vowel "a" and the following consonant "s", not shown.

Page 33, line 25 – page 34, line 25:

In accordance with the generation of the singing voices as described above, the singing voice control is carried out based on the information in the performance data S_1 to S_3 and the information in the transition track T_R. More specifically, before and after the time point T1, the tone generator control information corresponding to the state information of the transition [[sate]] state Attack in the track T_R and the information of the transition s_a in the track T_P are read out from the state transition DB 14c in FIG. 7 to control the tone generator circuit 28, whereby minute changes in pitch, amplitude, and the like are added to the singing voice "s_a".

The control time period at this time corresponds to the duration time designated by the state information of the attack transition state Attack. Further, before and after the time point T2, the tone generator control information corresponding to the state information of the note transition state NtN in the track T_R and the information of the transition a_i in the track T_P, and the pitch information D_3 in the performance data S_2 is read out from the DB 14c to control the tone generator circuit 28, whereby minute changes in pitch, amplitude, and the like are added to the singing voice "a_i". The control time period at this time corresponds to the duration time designated by the state information of the note transition state NtN. Further, immediately before the time point T4, the tone generator control information corresponding to the state information of the release transition state Release in the track T_R and the information of the vowel a in the track T_P, and the pitch information E_3 in the performance data S_3 is read out from the DB 14c to control the tone generator circuit 28, whereby minute changes in pitch, amplitude, and the like are added to the singing voice "a". The control time period at this time corresponds to the duration time designated by the state information of the release transition state Release. According to the singing voice control described above, it is possible to synthesize natural singing voices with the feelings of attack, note transition, and release.

Page 51, line 25 – page 52, line 15:

If the answer to the question of the step S168 is negative (N), in a step S174, a NONE transition time length corresponding to the steady portion of the preceding performance data (referred to as "pNONEs transition time length") is calculated. Since the reception of the present performance data has made definite the state of connection with the preceding performance data, the NONEs transition time length and the preceding release transition time length formed based

on the preceding performance data are discarded. More specifically, the assumption "silence is interposed between the present performance data and the next performance data" employed in the processing in a step S176, described hereinafter, is annuled. In the step S174, as shown in FIGS. 36A to 36C, in both of the cases of PhU State = Consonant Vowel and PhU State = Vowel, the pNONEs transition time length is calculated such that the boundary between T_1 and T_2 of the NtN transition time length from the preceding vowel coincides with the actual singing-starting time point (Current Note On) of the present performance data based on the actual singing-starting time point and the actual singing end time point of the preset performance data and the NtN transition time length. The FIG. 36A example differs from the FIG. 36B example in that the consonant singing length C is interposed in the consonant singing time.

Page 55, line 35 - page 56, line 25:

The phonetic unit connection pattern shown in FIG. 39A corresponds to a case of the preceding vowel "a" - silence - "sa". The silence singing length is calculated with the consonant singing length C being inserted to lengthen the consonant ("s" in this example) of a phonetic unit formed by a consonant and a vowel. The phonetic unit connection pattern shown in FIG. 39B corresponds to a case of the preceding vowel "a" - silence - "pa". The silence singing length is calculated without the consonant singing length being inserted for a phonetic unit formed by a consonant and a vowel. The phonetic unit connection pattern shown in FIG. 39C corresponds to a case of the preceding vowel "a" - silence - "na". The silence singing length is calculated with the consonant singing length C being inserted to lengthen the consonant ("n" in this example) of a phonetic unit formed by a consonant (nasal sound or half vowel) and a vowel. The phonetic unit connection pattern shown in FIG. 39D is the same as the FIG. 39C example except that the

consonant singing length C is not inserted. The phonetic unit connection pattern shown in FIG. 39E corresponds to a case of the preceding vowel "a" - silence - "i". The silence singing length is calculated for a phonetic unit formed by vowels alone (the same applies to a phonetic unit formed by consonants (nasal sounds) alone).

///

///

///

///

///

///

///

///

///

///

///

///

///

///

///

///

///

///

///